**1 Git Started**

The staging area is a file, generally contained in your Git directory, that stores information about what will go into your next commit. Its technical name in Git parlance is the “index”, but the phrase “staging area” works just as well.

Git comes with a tool called git config that lets you get and set configuration variables that control all aspects of how Git looks and operates. These variables can be stored in three different places

1. /etc/gitconfig file: Contains values applied to every user on the system and all their repositories. If you pass the option --system to git config, it reads and writes from this file specifically. (Because this is a system configuration file, you would need administrative or superuser privilege to make changes to it.)

2. ~/.gitconfig or ~/.config/git/config file: Values specific personally to you, the user. You can make Git read and write to this file specifically by passing the --global option, and this affects *all* of the repositories you work with on your system.

3. config file in the Git directory (that is, .git/config) of whatever repository you’re currently using: Specific to that single repository. You can force Git to read from and write to this file with the --local option, but that is in fact the default. (Unsurprisingly, you need to be located somewhere in a Git repository for this option to work properly.)

Each level overrides values in the previous level

On Windows systems, Git looks for the .gitconfig file in the $HOME directory (C:Users\$USER for most people). It also still looks for /etc/gitconfig, although it’s relative to the MSys root, which is wherever you decide to install Git on your Windows system when you run the installer. If you are using version 2.x or later of Git for Windows, there is also a system-level config file at C:\Documents and Settings\All Users\Application Data\Git\config on Windows XP, and in C:\ProgramData\Git\config on Windows Vista and newer. This config file can only be changed by git config -f <file> as an admin.

**Your Identity**

The first thing you should do when you install Git is to set your user name and email address. This is important because every Git commit uses this information, and it’s immutably baked into the commits you start creating:

$ git config --global user.name "John Doe"

$ git config --global user.email johndoe@example.com

**Your Editor**

Now that your identity is set up, you can configure the default text editor that will be used when Git needs you to type in a message. If not configured, Git uses your system’s default editor.If you want to use a different text editor, such as Emacs, you can do the following:

$ git config --global core.editor emacs

If you want to check your configuration settings, you can use the git config --list command to list all the settings Git can find at that point:

You can also check what Git thinks a specific key’s value is by typing git config <key>:

$ git config user.name

Since Git might read the same configuration variable value from more than one file, it’s possible that you have an unexpected value for one of these values and you don’t know why. In cases like that, you can query Git as to the *origin* for that value, and it will tell you which configuration file had the final say in setting that

value:

$ git config --show-origin rerere.autoUpdate

file:/home/johndoe/.gitconfig false

If you ever need help while using Git, there are two equivalent ways to get the comprehensive manual page (manpage) help for any of the Git commands:

$ git help <verb>

In addition, if you don’t need the full-blown manpage help, but just need a quick refresher on the available options for a Git command, you can ask for the more concise “help” output with the -h or --help options, as in:

$ git add –h

**2 Git Basics**

**Short Status**

While the git status output is pretty comprehensive, it’s also quite wordy. Git also has a short status flag so you can see your changes in a more compact way. If you run git status -s or git status --short you get a far more simplified output from the command:

$ git status –s

New files that aren’t tracked have a ?? next to them, new files that have been added to the staging area have an A, modified files have an M and so on. There are two columns to the output - the lefthand column indicates the status of the staging area and the right-hand column indicates the status of the working tree. So for example in that output, the README file is modified in the working directory but not yet staged, while the lib/simplegit.rb file is modified and staged. The Rakefile was modified, staged and then modified again, so there are changes to it that are both staged and unstaged.

**Ignoring Files**

Often, you’ll have a class of files that you don’t want Git to automatically add or even show you as being untracked. These are generally automatically generated files such as log files or files produced by your build system. In such cases, you can create a file listing patterns to match them named .gitignore. Here is an example .gitignore file:

$ cat .gitignore

\*.[oa]

\*~

The first line tells Git to ignore any files ending in “.o” or “.a” — object and archive files that may be the product of building your code. The second line tells Git to ignore all files whose names end with a tilde (~), which is used by many text editors such as Emacs to mark temporary files. You may also include a log, tmp, or pid directory; automatically generated documentation; and so on. Setting up a .gitignore file for your new repository before you get going is generally a good idea so you don’t accidentally commit files that you really don’t want in your Git repository. The rules for the patterns you can put in the .gitignore file are as follows:

* Blank lines or lines starting with # are ignored.
* Standard glob patterns work, and will be applied recursively throughout the entire working tree.
* You can start patterns with a forward slash (/) to avoid recursivity.
* You can end patterns with a forward slash (/) to specify a directory.
* You can negate a pattern by starting it with an exclamation point (!).

Glob patterns are like simplified regular expressions that shells use. An asterisk (\*) matches zero or more characters; [abc] matches any character inside the brackets (in this case a, b, or c); a question mark (?) matches a single character; and brackets enclosing characters separated by a hyphen ([0-9]) matches any character between them (in this case 0 through 9). You can also use two asterisks to match nested directories; a/\*\*/z would match a/z, a/b/z, a/b/c/z, and so on. Here is another example .gitignore file:

# ignore all .a files

\*.a

# but do track lib.a, even though you're ignoring .a files above

!lib.a

# only ignore the TODO file in the current directory, not subdir/TODO

/TODO

# ignore all files in any directory named build

build/

# ignore doc/notes.txt, but not doc/server/arch.txt

doc/\*.txt

# ignore all .pdf files in the doc/ directory and any of its subdirectories

doc/\*\*/\*.pdf

**Viewing Your Staged and Unstaged Changes**

To see what you’ve changed but not yet staged, type git diff with no other arguments. That command compares what is in your working directory with what is in your staging area. The result tells you the changes you’ve made that you haven’t yet staged.

If you want to see what you’ve staged that will go into your next commit, you can use git diff --staged. This command compares your staged changes to your last commit:

It’s important to note that git diff by itself doesn’t show all changes made since your last commit — only changes that are still unstaged. If you’ve staged all of your changes, git diff will give you no output.

Now you can use git diff to see what is still unstaged and git diff --cached to see what you’ve staged so far (--staged and --cached are synonyms):

*Git Diff in an External Tool*

We will continue to use the git diff command in various ways throughout the rest of the book. There is another way to look at these diffs if you prefer a graphical or external diff viewing program instead. If you run git difftool instead of git diff, you can view any of these diffs in software like emerge, vimdiff and many more (including commercial products). Run git difftool --tool-help to see what is available on your system

If you want to skip the staging area, Git provides a simple shortcut. Adding the -a option to the git commit command makes Git automatically stage every file that is already tracked before doing the commit,

$ git commit -am 'added new benchmarks'

To remove a file from Git, you have to remove it from your tracked files (more accurately, remove it from your staging area) and then commit. The git rm command does that, and also removes the file from your working directory so you don’t see it as an untracked file the next time around.

The next time you commit, the file will be gone and no longer tracked. If you modified the file and added it to the staging area already, you must force the removal with the -f option. This is a safety feature to prevent accidental removal of data that hasn’t yet been recorded in a snapshot and that can’t be recovered from Git.

Another useful thing you may want to do is to keep the file in your working tree but remove it from your staging area. In other words, you may want to keep the file on your hard drive but not have Git track it anymore. This is particularly useful if you forgot to add something to your .gitignore file and accidentally staged it, like a large log file or a bunch of .a compiled files. To do this, use the --cached option:

$ git rm --cached README

That means you can do things such as:

$ git rm log/\*.log

$ git rm /\*~

**Moving Files**

Unlike many other VCS systems, Git doesn’t explicitly track file movement. If you rename a file in Git, no metadata is stored in Git that tells it you renamed the file. However, Git is pretty smart about figuring that out after the fact — we’ll deal with detecting file movement a bit later. Thus it’s a bit confusing that Git has a mv command. If you want to rename a file in Git, you can run something like:

$ git mv file\_from file\_to

However, this is equivalent to running something like this:

$ mv README.md README

$ git rm README.md

$ git add README

Git figures out that it’s a rename implicitly, so it doesn’t matter if you rename a file that way or with the mv command. The only real difference is that git mv is one command instead of three — it’s a convenience function. More importantly, you can use any tool you like to rename a file, and address

the add/rm later, before you commit.

**Viewing the Commit History**

The most basic and powerful tool to do this is the git log command.

One of the more helpful options is -p or --patch, which shows the difference (the *patch* output) introduced in each commit. You can also limit the number of log entries displayed, such as using -2 to show only the last two entries.

$ git log -p -2

if you want to see some abbreviated stats for each commit, you can use the --stat option

As you can see, the --stat option prints below each commit entry a list of modified files, how many files were changed, and how many lines in those files were added and removed. It also puts a summary of the information at the end.

Another really useful option is --pretty. This option changes the log output to formats other than the default. A few prebuilt options are available for you to use. The oneline option prints each commit on a single line, which is useful if you’re looking at a lot of commits. In addition, the short, full, and fuller options show the output in roughly the same format but with less or more information, respectively:

$ git log --pretty=oneline

The most interesting option is format, which allows you to specify your own log output format. This is especially useful when you’re generating output for machine parsing — because you specify the format explicitly, you know it won’t change with updates to Git:

$ git log --pretty=format:"%h - %an, %ar : %s"

ca82a6d - Scott Chacon, 6 years ago : changed the version number

085bb3b - Scott Chacon, 6 years ago : removed unnecess

**Option Description of Output**

%H Commit hash

%h Abbreviated commit hash

%T Tree hash

%t Abbreviated tree hash

%P Parent hashes

%p Abbreviated parent hashes

%an Author name

%ae Author email

%ad Author date (format respects the --date=option)

%ar Author date, relative

%cn Committer name

%ce Committer email

%cd Committer date

%cr Committer date, relative

%s Subject

The oneline and format options are particularly useful with another log option called --graph. This option adds a nice little ASCII graph showing your branch and merge history:

*Table 2. Common options to* git log

**Option Description**

-p Show the patch introduced with each commit.

--stat Show statistics for files modified in each commit.

--shortstat Display only the changed/insertions/deletions line from the --stat command.

--name-only Show the list of files modified after the commit information.

--name-status Show the list of files affected with added/modified/deleted information as well.

--abbrev-commit Show only the first few characters of the SHA-1 checksum instead of all 40.

--relative-date Display the date in a relative format (for example, “2 weeks ago”) instead of using the full date format.

--graph Display an ASCII graph of the branch and merge history beside the log output.

--pretty Show commits in an alternate format. Options include oneline, short, full,fuller, and format (where you specify your own format).

--oneline Shorthand for --pretty=oneline --abbrev-commit used together.

**Limiting Log Output**

In addition to output-formatting options, git log takes a number of useful limiting options

* -<n>, where n is any integer to show the last n commits
* --since and --until the time-limiting options.

$ git log --since=2.weeks

This command works with lots of formats — you can specify a specific date like "2008-01-15", or a relative date such as "2 years 1 day 3 minutes ago"

* The --author option allows you to filter on a specific author,
* the --grep option lets you search for keywords in the commit messages

Another really helpful filter is the -S option (colloquially referred to as Git’s “pickaxe” option), which takes a string and shows only those commits that changed the number of occurrences of that string.

$ git log -S FindNextStructure

The last really useful option to pass to git log as a filter is a path. If you specify a directory or file name, you can limit the log output to commits that introduced a change to those files. This is always the last option and is generally preceded by double dashes (--) to separate the paths from the options.

**Option Description**

-<n> Show only the last n commits

--since, --after Limit the commits to those made after the specified date.

--until, --before Limit the commits to those made before the specified date.

--author Only show commits in which the author entry matches the specified string.

--committer Only show commits in which the committer entry matches the specified string.

--grep Only show commits with a commit message containing the string

-S Only show commits adding or removing code matching the string

$ git log --pretty="%h - %s" --author='Junio C Hamano' --since="2008-10-01"

--before="2008-11-01" --no-merges -- testData/

To prevent the display of merge commits cluttering up your log history, simply add the log option --no-merges

**Undoing Things**

If you want to redo that commit, make the additional changes you forgot, stage them, and commit again using the --amend option:

$ git commit –amend

**Unstaging a Staged File**

git reset HEAD <file> to unstagea file

**Unmodifying a Modified File**

$ git checkout -- CONTRIBUTING.md

It’s important to understand that git checkout -- <file> is a dangerous command. Any changes you made to that file are gone — Git just copied another file over it. Don’t ever use this command unless you absolutely know that you don’t want the file.

Remember, anything that is *committed* in Git can almost always be recovered. Even commits that were on branches that were deleted or commits that were overwritten with an --amend commit can be recovered (see Data Recovery for data recovery). However, anything you lose that was never committed is likely never to be seen again.

**Working with Remotes**

You can also specify -v, which shows you the URLs that Git has stored for the shortname to be used when reading and writing to that remote:

**Adding Remote Repositories**

To add a new remote Git repository as a shortname you can reference easily, run

$ git remote add <shortname> <url>:

**Fetching and Pulling from Your Remotes**

As you just saw, to get data from your remote projects, you can run:

$ git fetch <remote>

It’s important to note that the git fetch command only downloads the data to your local repository —

git pull command to automatically fetch and then merge that remote branch into your current branch

Running git pull generally fetches data from the server you originally cloned from and automatically tries to merge it into the code you’re currently working on.

**Pushing to Your Remotes**

The command for this is simple: git push <remote> <branch>. If you want to push your master branch to your origin server (again, cloning generally sets up both of those names for you automatically), then you can run this to push any commits you’ve done back up to the server:

$ git push origin master

**Inspecting a Remote**

If you want to see more information about a particular remote, you can use the git remote show <remote> command

$ git remote show origin

This command shows which branch is automatically pushed to when you run git push while on certain branches. It also shows you which remote branches on the server you don’t yet have, which remote branches you have that have been removed from the server, and multiple local branches that are able to merge automatically with their remote-tracking branch when you run git pull.

**Renaming and Removing Remotes**

For instance, if you want to rename pb to paul, you can do so with git remote rename

If you want to remove a remote for some reason — you’ve moved the server or are no longer using a particular mirror, or perhaps a contributor isn’t contributing anymore — you can either use git remote remove or git remote rm:

**Tagging**

**Listing Your Tags**

Listing the available tags in Git is straightforward. Just type git tag (with optional -l or --list):

$ git tag

You can also search for tags that match a particular pattern

$ git tag -l "v1.8.5\*"

If you want just the entire list of tags, running the command git tag implicitly assumes you want a listing and provides one; the use of -l or --list in this case is optional. If, however, you’re supplying a wildcard pattern to match tag names, the use of –l or --list is mandatory.

**Creating Tags**

Git supports two types of tags: *lightweight* and *annotated*.

A lightweight tag is very much like a branch that doesn’t change — it’s just a pointer to a specific commit.

Annotated tags, however, are stored as full objects in the Git database. The easiest way is to specify -a when you run the tag command:

$ git tag -a v1.4 -m "my version 1.4"

The -m specifies a tagging message, which is stored with the tag. If you don’t specify a message for an annotated tag, Git launches your editor so you can type it in.

You can see the tag data along with the commit that was tagged by using the git show command

$ git show v1.4

**Lightweight Tags**

Another way to tag commits is with a lightweight tag. To create a lightweight tag, just provide a tag name:

$ git tag v1.4

**Tagging Later**

To tag that commit, you specify the commit checksum (or part of it) at the end of the command:

$ git tag -a v1.2 9fceb02

**Sharing Tags**

By default, the git push command doesn’t transfer tags to remote servers. You will have to explicitly push tags to a shared server after you have created them. This process is just like sharing remote branches — you can run git push origin <tagname>.

$ git push origin v1.5

you can also use the --tags option to the git push command. This will transfer all of your tags to the remote server that are not already there.

$ git push origin –tags

**Checking out Tags**

If you want to view the versions of files a tag is pointing to, you can do a git checkout

$ git checkout 2.0.0

**Git Aliases**

If you don’t want to type the entire text of each of the Git commands, you can easily set up an alias for each command using git config

$ git config --global alias.unstage 'reset HEAD --'

It’s also common to add a last command, like this:

$ git config --global alias.last 'log -1 HEAD'

However, maybe you want to run an external command, rather than a Git subcommand. In that case, you start the command with a ! character. This is useful if you write your own tools that work with a Git repository. We can demonstrate by aliasing git visual to run gitk:

$ git config --global alias.visual '!gitk'

**3 Git Branching**

**4.Customizing Git**

**Git Configuration**

As you read briefly in Getting Started, you can specify Git configuration settings with the git config command. One of the first things you did was set up your name and email address:

$ git config --global user.name "John Doe"

$ git config --global user.email johndoe@example.com

First, a quick review: Git uses a series of configuration files to determine non-default behavior that you may want. The first place Git looks for these values is in the system-wide /etc/gitconfig file, which contains settings that are applied to every user on the system and all of their repositories. If you pass the option --system to git config, it reads and writes from this file specifically.

The next place Git looks is the ~/.gitconfig (or ~/.config/git/config) file, which is specific to each user. You can make Git read and write to this file by passing the --global option.

Finally, Git looks for configuration values in the configuration file in the Git directory (.git/config) of whatever repository you’re currently using. These values are specific to that single repository, and represent passing the --local option to git config. (If you don’t specify which level you want to work with, this is the default.)

Each of these “levels” (system, global, local) overwrites values in the previous level, so values in .git/config trump those in /etc/gitconfig, for instance

core.editor:

By default, Git uses whatever you’ve set as your default text editor via one of the shell environment variables VISUAL or EDITOR, or else falls back to the vi editor to create and edit your commit and tag messages. To change that default to something else, you can use the core.editor setting:

$ git config --global core.editor emacs

commit.template:

If you set this to the path of a file on your system, Git will use that file as the default initial message when you commit.

$ git config --global commit.template ~/.gitmessage.txt

core.autocrlf

$ git config --global core.autocrlf true

$ git config --global core.autocrlf input

If you’re a Windows programmer doing a Windows-only project, then you can turn off this functionality, recording the carriage returns in the repository by setting the config value to false:

$ git config --global core.autocrlf false

**5 Git Internals**

**Plumbing and Porcelain**

When you run git init in a new or existing directory, Git creates the .git directory, which is where almost everything that Git stores and manipulates is located.

a newlyinitialized .git directory typically looks like:

$ ll -la .git

total 11

drwxr-xr-x 1 asaki 1049089 0 Sep 27 18:37 ./

drwxr-xr-x 1 asaki 1049089 0 Sep 27 18:37 ../

-rw-r--r-- 1 asaki 1049089 130 Sep 27 18:37 config

-rw-r--r-- 1 asaki 1049089 73 Sep 27 18:37 description

-rw-r--r-- 1 asaki 1049089 23 Sep 27 18:37 HEAD

drwxr-xr-x 1 asaki 1049089 0 Sep 27 18:37 hooks/

drwxr-xr-x 1 asaki 1049089 0 Sep 27 18:37 info/

drwxr-xr-x 1 asaki 1049089 0 Sep 27 18:37 objects/

drwxr-xr-x 1 asaki 1049089 0 Sep 27 18:37 refs/

$ ls -F1

config

description

HEAD

hooks/

info/

objects/

refs/

Depending on your version of Git, you may see some additional content there, but this is a fresh git init repository — it’s what you see by default. The description file is used only by the GitWeb program, so don’t worry about it. The config file contains your project-specific configuration options, and the info directory keeps a global exclude file for ignored patterns that you don’t want to track in a .gitignore file. The hooks directory contains your client- or server-side hook scripts, which are discussed in detail in Git Hooks.

This leaves four important entries: the HEAD and (yet to be created) index files, and the objects and refs directories. These are the core parts of Git. The objects directory stores all the content for your database, the refs directory stores pointers into commit objects in that data (branches, tags,

remotes and more), the HEAD file points to the branch you currently have checked out, and the index file is where Git stores your staging area information. You’ll now look at each of these sections in detail to see how Git operates.

**Git Objects**

Git is a content-addressable filesystem. Great. What does that mean? It means that at the core of Git is a simple key-value data store. What this means is that you can insert any kind of content into a Git repository, for which Git will hand you back a unique key you can use later to retrieve that content.

$ find .git/objects

.git/objects

.git/objects/info

.git/objects/pack

$ find .git/objects -type f

Git has initialized the objects directory and created pack and info subdirectories in it, but there are no regular files. Now, let’s use git hash-object to create a new data object and manually store it in your new Git database

$ echo 'test content' | git hash-object -w –stdin

In its simplest form, git hash-object would take the content you handed to it and merely return the unique key that *would* be used to store it in your Git database. The -w option then tells the command to not simply return the key, but to write that object to the database. Finally, the --stdin option tells git hash-object to get the content to be processed from stdin; otherwise, the command would expect a filename argument at the end of the command containing the content to be used.

The output from the above command is a 40-character checksum hash. This is the SHA-1 hash — a checksum of the content you’re storing plus a header, which you’ll learn about in a bit. Now you can see how Git has stored your data:

$ find .git/objects -type f

.git/objects/d6/70460b4b4aece5915caf5c68d12f560a9fe3e4

This is how Git stores the content initially — as a single file per piece of content, named with the SHA-1 checksum of the content and its header. The subdirectory is named with the first 2 characters of the SHA-1, and the filename is the remaining 38 characters.

Once you have content in your object database, you can examine that content with the git cat-file command. This command is sort of a Swiss army knife for inspecting Git objects. Passing -p to catfile instructs the command to first figure out the type of content, then display it appropriately:

$ git cat-file -p d670460b4b4aece5915caf5c68d12f560a9fe3e4

test content

Now, you can add content to Git and pull it back out again. You can also do this with content in files. For example, you can do some simple version control on a file. First, create a new file and save its contents in your database:

$ echo 'version 1' > test.txt

$ git hash-object -w test.txt

83baae61804e65cc73a7201a7252750c76066a30

Then, write some new content to the file, and save it again:

$ echo 'version 2' > test.txt

$ git hash-object -w test.txt

1f7a7a472abf3dd9643fd615f6da379c4acb3e3a

Your object database now contains both versions of this new file (as well as the first content you stored there):

$ find .git/objects -type f

.git/objects/1f/7a7a472abf3dd9643fd615f6da379c4acb3e3a

.git/objects/83/baae61804e65cc73a7201a7252750c76066a30

.git/objects/d6/70460b4b4aece5915caf5c68d12f560a9fe3e4

You can have Git tell you the object type of any object in Git, given its SHA-1 key, with git cat-file -t:

$ git cat-file -t 1f7a7a472abf3dd9643fd615f6da379c4acb3e3a

Blob

**Tree Objects**

The next type of Git object we’ll examine is the *tree*, which solves the problem of storing the filename and also allows you to store a group of files together. Git stores content in a manner similar to a UNIX filesystem, but a bit simplified. All the content is stored as tree and blob objects, with trees corresponding to UNIX directory entries and blobs corresponding more or less to inodes or file contents. A single tree object contains one or more entries, each of which is the SHA-1 hash of a blob or subtree with its associated mode, type, and filename. For example, the most recent tree in a project may look something like this:

$ git cat-file -p master^{tree}

100644 blob a906cb2a4a904a152e80877d4088654daad0c859 README

100644 blob 8f94139338f9404f26296befa88755fc2598c289 Rakefile

040000 tree 99f1a6d12cb4b6f19c8655fca46c3ecf317074e0 lib

The master^{tree} syntax specifies the tree object that is pointed to by the last commit on your master branch. Notice that the lib subdirectory isn’t a blob but a pointer to another tree:

**Commit Objects**

This is essentially what Git does when you run the git add and git commit

commands — it stores blobs for the files that have changed, updates the index, writes out trees, and writes commit objects that reference the top-level trees and the commits that came immediately before them. These three main Git objects — the blob, the tree, and the commit — are initially stored as separate files in your .git/objects directory. Here are all the objects in the example directory now, commented with what they store:

**Object Storage**

**Git References**

Instead, itwould be easier if you had a file in which you could store that SHA-1 value under a simple name so you could use that simple name rather than the raw SHA-1 value.

In Git, these simple names are called “references” or “refs”; you can find the files that contain those SHA-1 values in the .git/refs directory. In the current project, this directory contains no files, but it does contain a simple structure:

$ find .git/refs

.git/refs

.git/refs/heads

.git/refs/tags

$ find .git/refs -type f

That’s basically what a branch in Git is: a simple pointer or reference to the head of a line of work.

**The HEAD**

The HEAD file is a symbolic reference to the branch you’re currently on. By symbolic reference, we mean that unlike a normal reference, it doesn’t generally contain a SHA-1 value but rather a pointer to another reference.

$ cat .git/HEAD

ref: refs/heads/master

When you run git commit, it creates the commit object, specifying the parent of that commit object to be whatever SHA-1 value the reference in HEAD points to.

**Tags**

We just finished discussing Git’s three main object types (*blobs*, *trees* and *commits*), but there is a fourth. The *tag* object is very much like a commit object — it contains a tagger, a date, a message, and a pointer. The main difference is that a tag object generally points to a commit rather than a tree. It’s like a branch reference, but it never moves — it always points to the same commit but gives it a friendlier name

That is all a lightweight tag is — a reference that never moves. An annotated tag is more complex, however. If you create an annotated tag, Git creates a tag object and then writes a reference to point to it rather than directly to the commit. You can see this by creating an annotated tag (using the -a option):

$ git tag -a v1.1 1a410efbd13591db07496601ebc7a059dd55cfe9 -m 'test tag'

Here’s the object SHA-1 value it created:

$ cat .git/refs/tags/v1.1

9585191f37f7b0fb9444f35a9bf50de191beadc2

**Remotes**

The third type of reference that you’ll see is a remote reference. If you add a remote and push to it, Git stores the value you last pushed to that remote for each branch in the refs/remotes directory

Then, you can see what the master branch on the origin remote was the last time you

communicated with the server, by checking the refs/remotes/origin/master file:

$ cat .git/refs/remotes/origin/master

ca82a6dff817ec66f44342007202690a93763949

Remote references differ from branches (refs/heads references) mainly in that they’re considered read-only. You can git checkout to one, but Git won’t point HEAD at one, so you’ll never update it with a commit command. Git manages them as bookmarks to the last known state of where those branches were on those servers.

**Packfiles**

The initial format in which Git saves objects on disk is called a “loose” object

format. However, occasionally Git packs up several of these objects into a single binary file called a “packfile” in order to save space and be more efficient. Git does this if you have too many loose objects around, if you run the git gc command manually

If you look in your objects directory, you’ll find that most of your objects are gone, and a new pair of files has appeared:

$ find .git/objects -type f

.git/objects/bd/9dbf5aae1a3862dd1526723246b20206e5fc37

.git/objects/d6/70460b4b4aece5915caf5c68d12f560a9fe3e4

.git/objects/info/packs

.git/objects/pack/pack-978e03944f5c581011e6998cd0e9e30000905586.idx

.git/objects/pack/pack-978e03944f5c581011e6998cd0e9e30000905586.pack

The objects that remain are the blobs that aren’t pointed to by any commit — in this case, the “what is up, doc?” example and the “test content” example blobs you created earlier. Because you never added them to any commits, they’re considered dangling and aren’t packed up in your new packfile.

The other files are your new packfile and an index. The packfile is a single file containing the contents of all the objects that were removed from your filesystem. The index is a file that contains offsets into that packfile so you can quickly seek to a specific object. What is cool is that although the objects on disk before you ran the gc command were collectively about 15K in size, the new packfile is only 7K. You’ve cut your disk usage by half by packing your objects.

How does Git do this? When Git packs objects, it looks for files that are named and sized similarly, and stores just the deltas from one version of the file to the next. You can look into the packfile and see what Git did to save space. The git verify-pack plumbing command allows you to see what was packed up:

$ git verify-pack -v .git/objects/pack/pack-978e03944f5c581011e6998cd0e9e30000905586.idx

**The Refspec**

$ git remote add origin https://github.com/schacon/simplegit-progit

Running the command above adds a section to your repository’s .git/config file, specifying the name of the remote (origin), the URL of the remote repository, and the *refspec* to be used for fetching:

[remote "origin"]

url = https://github.com/schacon/simplegit-progit

fetch = +refs/heads/\*:refs/remotes/origin/\*

The format of the refspec is, first, an optional +, followed by <src>:<dst>, where <src> is the pattern for references on the remote side and <dst> is where those references will be tracked locally. The + tells Git to update the reference even if it isn’t a fast-forward.In the default case that is automatically written by a git remote add command, Git fetches all the references under refs/heads/ on the server and writes them to refs/remotes/origin/ locally. So, if there is a master branch on the server, you can access the log of that branch locally via any of the following:

$ git log origin/master

$ git log remotes/origin/master

$ git log refs/remotes/origin/master

They’re all equivalent, because Git expands each of them to refs/remotes/origin/master.

If you want Git instead to pull down only the master branch each time, and not every other branch on the remote server, you can change the fetch line to refer to that branch only:

fetch = +refs/heads/master:refs/remotes/origin/master

This is just the default refspec for git fetch for that remote. If you want to do a one-time only fetch, you can specify the specific refspec on the command line, too. To pull the master branch on the remote down to origin/mymaster locally, you can run:

$ git fetch origin master:refs/remotes/origin/mymaster

You can also specify multiple refspecs for fetching in your configuration file. If you want to always fetch the master and experiment branches from the origin remote, add two lines:

[remote "origin"]

url = https://github.com/schacon/simplegit-progit

fetch = +refs/heads/master:refs/remotes/origin/master

fetch = +refs/heads/experiment:refs/remotes/origin/experiment

However, you can use namespaces (or directories) to accomplish something like that. If you have a QA team that pushes a series of branches, and you want to get the master branch and any of the QA team’s branches but nothing else, you can use a config section like this:

[remote "origin"]

url = https://github.com/schacon/simplegit-progit

fetch = +refs/heads/master:refs/remotes/origin/master

fetch = +refs/heads/qa/\*:refs/remotes/origin/qa/\*

If you have a complex workflow process that has a QA team pushing branches, developers pushing branches, and integration teams pushing and collaborating on remote branches, you can namespace them easily this way.

**Pushing Refspecs**

It’s nice that you can fetch namespaced references that way, but how does the QA team get their branches into a qa/ namespace in the first place? You accomplish that by using refspecs to push. If the QA team wants to push their master branch to qa/master on the remote server, they can run

$ git push origin master:refs/heads/qa/master

If they want Git to do that automatically each time they run git push origin, they can add a push value to their config file:

[remote "origin"]

url = https://github.com/schacon/simplegit-progit

fetch = +refs/heads/\*:refs/remotes/origin/\*

push = refs/heads/master:refs/heads/qa/master

Again, this will cause a git push origin to push the local master branch to the remote qa/master branch by default.

**Deleting References**

You can also use the refspec to delete references from the remote server by running something like this:

$ git push origin :topic

Because the refspec is <src>:<dst>, by leaving off the <src> part, this basically says to make the topic branch on the remote nothing, which deletes it.

Or you can use the newer syntax (available since Git v1.7.0):

$ git push origin --delete topic

**Maintenance and Data Recovery**

Occasionally, you may have to do some cleanup – make a repository more compact, clean up an imported repository, or recover lost work. This section will cover some of these scenarios.

**Maintenance**

Occasionally, Git automatically runs a command called “auto gc”. Most of the time, this command does nothing. However, if there are too many loose objects (objects not in a packfile) or too many packfiles, Git launches a full-fledged git gc command. The “gc” stands for garbage collect, and the command does a number of things: it gathers up all the loose objects and places them in packfiles, it consolidates packfiles into one big packfile, and it removes objects that aren’t reachable from any commit and are a few months old.

$ git gc

**Data Recovery**

Often, the quickest way is to use a tool called git reflog. As you’re working, Git silently records what your HEAD is every time you change it. Each time you commit or change branches, the reflog is updated. The reflog is also updated by the git update-ref command, which is another reason to use it instead of just writing the SHA-1 value to your ref files, as we covered in Git References. You can see where you’ve been at any time by running git reflog:

$ git reflog

1a410ef HEAD@{0}: reset: moving to 1a410ef

ab1afef HEAD@{1}: commit: modified repo.rb a bit

484a592 HEAD@{2}: commit: added repo.rb

To see the same information in a much more useful way, we can run git log -g, which will give you a normal log output for your reflog

Because the reflog data is kept in the .git/logs/ directory, you effectively have no reflog. How can you recover that commit at this point? One way is to use the git fsck utility, which checks your database for integrity. If you run it with the --full option, it shows you all objects that aren’t pointed to by another object:

$ git fsck --full

Checking object directories: 100% (256/256), done.

Checking objects: 100% (18/18), done.

dangling blob d670460b4b4aece5915caf5c68d12f560a9fe3e4

dangling commit ab1afef80fac8e34258ff41fc1b867c702daa24b

dangling tree aea790b9a58f6cf6f2804eeac9f0abbe9631e4c9

dangling blob 7108f7ecb345ee9d0084193f147cdad4d2998293

Vous pouvez executer la commande count-objects pour voir rapidement combien d’espace disque vous utilisez :

$ git count-objects -v

**Appendix A: Git in Other Environments**

gitk is a graphical history viewer. Think of it like a powerful GUI shell over git log and git grep

Gitk is easiest to invoke from the command-line. Just cd into a Git repository, and type:

$ gitk [git log options]

Gitk accepts many command-line options, most of which are passed through to the underlying git log action. Probably one of the most useful is the --all flag, which tells gitk to show commits reachable from *any* ref, not just HEAD.

git-gui, on the other hand, is primarily a tool for crafting commits. It, too, is easiest to invoke from the command line:

$ git gui

gitk and git-gui are examples of task-oriented tools. Each of them is tailored for a specific purpose (viewing history and creating commits, respectively), and omit the features not necessary for that task.

**Appendix C: Git Commands**

**Setup and Config**

**git config**

Git has a default way of doing hundreds of things. For a lot of these things, you can tell Git to default to doing them a different way, or set your preferences.

**git help**

The git help command is used to show you all the documentation shipped with Git about any command

$ git help <command>

**Getting and Creating Projects**

**git init**

To take a directory and turn it into a new Git repository so you can start version controlling it, you can simply run git init.

**git clone**

The git clone command is actually something of a wrapper around several other commands. It creates a new directory, goes into it and runs git init to make it an empty Git repository, adds a remote (git remote add) to the URL that you pass it (by default named origin), runs a git fetch from that remote repository and then checks out the latest commit into your working directory with git checkout.

the --bare option to create a copy of a Git repository with no working directory

the --recurse-submodules option to makecloning a repository with submodules a little simpler.

**Basic Snapshotting:**

For the basic workflow of staging content and committing it to your history, there are only a few basic commands.

**git add**

The git add command adds content from the working directory into the staging area (or “index”) for the next commit.

**git status**

The git status command will show you the different states of files in your working directory and staging area. Which files are modified and unstaged and which are staged but not yet committed

**git diff:**

The git diff command is used when you want to see differences between any two trees. This could be the difference between your working environment and your staging area (git diff by itself), between your staging area and your last commit (git diff --staged), or between two commits (git diff master branchB)

We use it to look for possible whitespace issues before committing with the --check option in Commit Guidelines.

We see how to check the differences between branches more effectively with the git diff A...B syntax in Determining What Is Introduced.

We use it to filter out whitespace differences with -b and how to compare different stages of conflicted files with --theirs, --ours and --base in Advanced Merging.

Finally, we use it to effectively compare submodule changes with --submodule in Starting with Submodules.

**git difftool:**

The git difftool command simply launches an external tool to show you the difference between two trees in case you want to use something other than the built in git diff command.

**git commit**

The git commit command takes all the file contents that have been staged with git add and records a new permanent snapshot in the database and then moves the branch pointer on the current branch up to it.

We first cover the basics of committing in Committing Your Changes. There we also demonstrate how to use the -a flag to skip the git add step in daily workflows and how to use the -m flag to pass a commit message in on the command line instead of firing up an editor

the --amend option to redo the most recent commit.

the -S flag sign commits cryptographically

**git reset:**

The git reset command is primarily used to undo things, as you can possibly tell by the verb. It moves around the HEAD pointer and optionally changes the index or staging area and can also

optionally change the working directory if you use --hard. This final option makes it possible for this command to lose your work if used incorrectly, so make sure you understand it before using it.

We first effectively cover the simplest use of git reset in Unstaging a Staged File, where we use it to unstage a file we had run git add on.

**git rm**

The git rm command is used to remove files from the staging area and working directory for Git

including recursively removing files and only removing files from the staging area but leaving them in the working directory with --cached.

**git mv**

The git mv command is a thin convenience command to move a file and then run git add on the new file and git rm on the old file.

**git clear**

The git mv command is a thin convenience command to move a file and then run git add on the new file and git rm on the old file.

**Branching and Merging**

There are just a handful of commands that implement most of the branching and merging

functionality in Git.

**git branch**

The git branch command is actually something of a branch management tool. It can list the branches you have, create a new branch, delete branches and rename branches

In Tracking Branches we use the git branch -u option to set up a tracking branch.

**git checkout**

The git checkout command is used to switch branches and check content out into your working directory.

We see how to use it to start tracking branches with the --track flag in Tracking Branches.

**git merge**

The git merge tool is used to merge one or more branches into the branch you have checked out. It will then advance the current branch to the result of the merge.

We use it to reintroduce file conflicts with --conflict=diff3 in Checking Out Conflicts.

We go into closer detail on its relationship with git reset in Reset Demystified.

**git mergetool**

The git mergetool command simply launches an external merge helper in case you have issues with a merge in Git.

**git log**

The git log command is used to show the reachable recorded history of a project from the most recent commit snapshot backwards.

There we look at the -p and --stat option to get an idea of what was introduced in each commit and the --pretty and --oneline options to view the history more concisely, along with some simple date and author filtering options.

we use it with the --decorate option to easily visualize where our branch

pointers are located and we also use the --graph option to see what divergent histories look like

In Private Small Team and Commit Ranges we cover the branchA..branchB syntax to use the git log command to see what commits are unique to a branch relative to another branch. In Commit Ranges we go through this fairly extensively

In Merge Log and Triple Dot we cover using the branchA...branchB format and the --left-right syntax to see what is in one branch or the other but not in both. In Merge Log we also look at how to use the --merge option to help with merge conflict debugging as well as using the --cc option to look at merge commit conflicts in your history.

In RefLog Shortnames we use the -g option to view the Git reflog through this tool instead of doing branch traversal.

In Searching we look at using the -S and -L options to do fairly sophisticated searches for something that happened historically in the code such as seeing the history of a function.

In Signing Commits we see how to use --show-signature to add a validation string to each commit in the git log output based on if it was validly signed or not.

**git stash**

The git stash command is used to temporarily store uncommitted work in order to clean out your working directory without having to commit unfinished work on a branch

**git tag**

The git tag command is used to give a permanent bookmark to a specific point in the code history. Generally this is used for things like releases. This command is introduced and covered in detail

**Sharing and Updating Projects**

**git fetch**

The git fetch command communicates with a remote repository and fetches down all the information that is in that repository that is not in your current one and stores it in your local database.

**git pull**

The git pull command is basically a combination of the git fetch and git merge commands, where Git will fetch from the remote you specify and then immediately try to merge it into the branch you’re on.

**git push**

The git push command is used to communicate with another repository, calculate what your local database has that the remote one does not, and then pushes the difference into the other repository. It requires write access to the other repository and so normally is authenticated somehow.

In Deleting Remote Branches we use the --delete flag to delete a branch on the server with git push.

We see how to use it to share tags that you have made with the --tags option in Sharing Tags.

In Publishing Submodule Changes we use the --recurse-submodules option to check that all of our submodules work has been published before pushing the superproject, which can be really helpful when using submodules.

**git remote**

The git remote command is a management tool for your record of remote repositories.

It is used in nearly every subsequent chapter in the book too, but always in the standard git remote add <name> <url> format

**git archive**

The git archive command is used to create an archive file of a specific snapshot of the project.

**git submodule**

The git submodule command is used to manage external repositories within a normal repositories. This could be for libraries or other types of shared resources. The submodule command has several sub-commands (add, update, sync, etc) for managing these resources.

**Inspection and Comparison**

**git show**

The git show command can show a Git object in a simple and human readable way. Normally you would use this to show the information about a tag or a commit.

**git shortlog**

The git shortlog command is used to summarize the output of git log. It will take many of the same options that the git log command will but instead of listing out all of the commits it will present a summary of the commits grouped by author.

**git describe**

The git describe command is used to take anything that resolves to a commit and produces a string that is somewhat human-readable and will not change. It’s a way to get a description of a commit that is as unambiguous as a commit SHA-1 but more understandable.We use git describe in Generating a Build Number and Preparing a Release to get a string to name our release file after.

**Debogage**

**git bisect**

The git bisect tool is an incredibly helpful debugging tool used to find which specific commit was the first one to introduce a bug or problem by doing an automatic binary search.

**git blame:**

The git blame command annotates the lines of any file with which commit was the last one to introduce a change to each line of the file and what person authored that commit. This is helpful in order to find the person to ask for more information about a specific section of your code.

**git grep**

The git grep command can help you find any string or regular expression in any of the files in your source code, even older versions of your project.

**Patching**

**git cherry-pick**

The git cherry-pick command is used to take the change introduced in a single Git commit and try to re-introduce it as a new commit on the branch you’re currently on. This can be useful to only take one or two commits from a branch individually rather than merging in the branch which takes all the changes.

**git rebase**

The git rebase command is basically an automated cherry-pick. It determines a series of commits and then cherry-picks them one by one in the same order somewhere else.

We also use it in an interactive scripting mode with the -i option

**git revert**

The git revert command is essentially a reverse git cherry-pick. It creates a new commit that applies the exact opposite of the change introduced in the commit you’re targeting, essentially undoing or reverting it.

**External Systems**

Git comes with a few commands to integrate with other version control systems.

**git svn**

The git svn command is used to communicate with the Subversion version control system as a client. This means you can use Git to checkout from and commit to a Subversion server.

**Administration**

**git gc**

The git gc command runs “garbage collection” on your repository, removing unnecessary files in your database and packing up the remaining files into a more efficient format.

**git fsck**

The git fsck command is used to check the internal database for problems or inconsistencies.

**git reflog**

The git reflog command goes through a log of where all the heads of your branches have been as you work to find commits you may have lost through rewriting histories.

We cover this command mainly in RefLog Shortnames, where we show normal usage to and how to use git log -g to view the same information with git log output.

We also go through a practical example of recovering such a lost branch in Data Recovery

**git filter-branch**

The git filter-branch command is used to rewrite loads of commits according to certain patterns, like removing a file everywhere or filtering the entire repository down to a single subdirectory for extracting a project

**Plumbing Commands**

ls-remote to look at the raw references on the server

ls-files to take a more raw look at what your staging area looks like.

rev-parse to take just about any string and turn it into an object SHA-1.

**Git distribue**

lancez donc git log --no-merges dessus pour voir a quoi ressemble un historique de *commits* avec des messages bien formates.

La difference principale reste que les fusions ont lieu du cote client plutot que

sur le serveur au moment de valider

Il est tres important de comprendre ceci si vous avez deja utilise Subversion, parce qu’il faut remarquer que les deux developpeurs n’ont pas modifie le meme fichier. Quand des fichiers differents ont été modifies, Subversion realise cette fusion automatiquement sur le serveur alors que Git necessite une fusion des modifications locale. John doit recuperer les modifications de Jessica et les fusionner avant d’etre autorise a pousser :

$ git log --no-merges issue54..origin/master

La syntaxe issue54..origin/master est un filtre du journal qui ordonne a Git de ne montrer que la liste des *commits* qui sont sur la seconde branche (dans ce cas origin/master) et qui ne sont pas sur la premiere (dans ce cas prob54). Nous aborderons cette syntaxe en detail dans Plages de *commits.*

Mais il y a un petit probleme — elle doit pousser son travail fusionne dans sa branche fonctionB sur la branche fonctionBee du serveur. Elle peut le faire en specifiant la branche locale suivie de deux points (:) suivi de la branche distante a la commande git push :

$ git push -u origin fonctionB:fonctionBee

Notez l’option -u. C’est un raccourci pour --set-upstream, qui configure les branches pour faciliter les poussees et les tirages plus tard.

Vous demarrez une nouvelle branche a partir de la branche origin/master

courante, y collez les modifications de fonctionB en resolvant les conflits, changez l’implementation et poussez le tout en tant que nouvelle branche

$ git checkout -b fonctionBv2 origin/master

$ git merge --no-commit --squash fonctionB

# (changement d'implementation)

$ git commit

$ git push macopie fonctionBv2

.

Pro Git (Vf)

$ git log –abrev-commit –pretty=oneline

Pour connaitre l’empreinte SHA sur laquelle pointe une branche

$ git rev-parse topic1

Git maintient en arriere-plan un historique des references ou sont passes HEAD et vos branches sur les derniers mois — ceci s’appelle le *reflog*. Vous pouvez le consulter avec la commande git reflog :

Si vous souhaitez consulter le n-ieme antecedent de votre HEAD, vous pouvez utiliser la reference @{n} du reflog, 5 dans cet exemple :

$ git show HEAD@{5}

Pour consulter le reflog au format git log, executez: git log -g. Veuillez noter que le reflog ne stocke que des informations locales, c’est un historique de ce que

vous avez fait dans votre depot

Alors, vous pouvez consulter le *commit* precedent en specifiant HEAD^, ce qui signifie ≪ le parent de HEAD ≫:

Vous pouvez egalement specifier un nombre apres ^ — par exemple, d921970^2 signifie ≪ le second parent de d921970 ≫. Cette syntaxe ne sert que pour les *commits* de fusion, qui ont plus d’un parent. Le premier parent est la branche depuis laquelle vous avez fusionne, et le second est le *commit* de la branche que vous avez fusionnee :

Une autre solution courante pour specifier une reference ancetre est le ~. Il fait egalement reference au premier parent, donc HEAD~ et HEAD^ sont equivalents. La difference apparait si vous specifiez un nombre. HEAD~2 signifie ≪ le premier parent du premier parent ≫, ou bien ≪ le grandparent ≫ ; on remonte les premiers parents autant de fois que demande.

HEAD~3 peut aussi s’ecrire HEAD^^^.

Vous pouvez egalement combiner ces syntaxes HEAD~3^2,

**Plages de** *commits*

**Double point**

Si vous voulez savoir ce qui n’a pas encore ete fusionne sur votre branche master depuis votre branche experiment, vous pouvez demander a Git de vous montrer un journal des *commits* avec master..experiment

$ git log origin/master..HEAD

Cette commande vous affiche tous les *commits* de votre branche courante qui ne sont pas sur la branche master du depot distant origin.

Vous pouvez egalement laisser tomber une borne de la syntaxe pour faire comprendre a Git que vous parlez de HEAD. Par exemple, vous pouvez obtenir les memes resultats que precedemment en tapant git log origin/master.. — Git utilise HEAD si une des bornes est manquante.

**Emplacements multiples**

Git vous permet cela avec ^ ou --not en prefixe de toute reference de laquelle vous ne souhaitez pas voir les *commits*. Les 3 commandes ci-apres sont equivalentes :

$ git log refA..refB

$ git log ^refA refB

$ git log refB --not refA

C’est utile car cela vous permet de specifier plus de 2 references dans votre requete, ce que vous ne pouvez accomplir avec la syntaxe double-point. Par exemple, si vous souhaitez voir les *commits* qui sont accessibles depuis refA et refB mais pas depuis refC, vous pouvez taper ces 2 commandes :

$ git log refA refB ^refC

$ git log refA refB --not refC

**Triple point**

La derniere syntaxe majeure de selection de plage de *commits* est la syntaxe triple-point qui specifie tous les *commits* accessibles par l’une des deux references, exclusivement

$ git log master...experiment

Une option courante a utiliser avec la commande log dans ce cas est --left-right qui vous montre la borne de la plage a laquelle ce *commit* appartient. Cela rend les donnees plus utiles :

$ git log --left-right master...experiment

Mode patch

$ git add –p

$ git reset –p

$ git checkout –p

$ git stash save -p

Remiser votre travail

$ git stash save

$ git stash list

$ git stash apply

vous voulez appliquer une remise plus ancienne, vous pouvez la specifier en la nommant, comme ceci : git stash apply stash@{2}. Si vous ne specifiez pas une remise, Git presume que vous voulez la remise la plus recente et essaye de l’appliquer

Par defaut, les modifications de vos fichiers sont re-appliquees, mais pas les indexations. Pour cela, vous devez executer la commande git stash apply avec l’option --index pour demander a Git d’essayer de re-appliquer les modifications de votre index.

Pour la supprimer, vous pouvez executer git stash drop avec le nom de la remise a supprimer

Vous pouvez egalement executer git stash pop pour appliquer et supprimer immediatement la remise de votre pile.

Il existe des variantes de remisages qui peuvent s’averer utiles.La premiere option

est l’option --keep-index de la commande stash save. Elle indique a Git de ne pas remiser ce qui aurait ete deja indexe au moyen de la commande git add.

Une autre option utile de stash est la possibilite de remiser les fichiers non suivis aussi bien que les fichiers suivis. Par defaut, git stash ne sauve que les fichiers qui sont deja suivis ou indexes. Si vous specifiez l’option --include-untracked ou -u, Git remisera aussi les fichiers non-suivis du repertoire de travail

Une option plus securisee consiste a lancer git stash --all pour tout sauvegarder dans une remise.

**Creer une branche depuis une remise**

vous pouvez executer git stash branch qui creera une nouvelle branche a votre place, recuperant le *commit* ou vous etiez lorsque vous avez cree la remise, re-appliquera votre travail dedans, et supprimera finalement votre remise si cela a reussi

$ git stash branch testchanges

**Nettoyer son repertoire de travail**

Pour supprimer tous les fichiers non-suivis, vous pouvez lancer git clean -f -d, qui effacera aussi tout sous-repertoire vide. L’option -f signifie ≪ force ≫, soit ≪ fais-le reellement ≫.

$ git clean -d -n

Si vous souhaitez visualiser ce qui serait fait, vous pouvez lancer la commande avec l’option -n qui signifie ≪ fais-le a blanc et montre-moi ce qui *serait* supprime ≫.

$ git clean -d –n

Par defaut, la commande git clean ne va supprimer que les fichiers non-suivis qui ne sont pas ignores. Tout fichier qui correspond a un motif de votre fichier .gitignore ou tout autre fichier similaire ne sera pas supprime. Si vous souhaitez supprimer aussi ces fichiers, comme par exemple les fichiers .o genere par un compilateur pour faire une compilation totale, vous pouvez ajouter l’option -x a la commande de nettoyage

Si vous ne savez pas ce que la commande git clean va effectivement supprimer, lancez-la une premiere fois avec -n par securite avant de transformer le -n en -f et nettoyer definitivement. Un autre choix pour s’assurer de ce qui va etre efface consiste a lancer la commande avec l’option –i ou --interactive.

La commande sera lancee en mode interactif.

$ git clean -x –i

**Recherche**

Git fournit quelques outils permettant rapidement de rechercher dans le code et les *commits* stockes dans votre base de données

**Git grep**

grep permet de rechercher facilement une chaine de caracteres ou une expression reguliere dans une arborescence validee ou dans le repertoire de travail. Pour tous les exemples qui suivent, nous allons utiliser le depot de Git lui-meme. Par defaut, git grep recherche dans le repertoire de travail. Vous pouvez passer l’option -n pour

afficher les numeros des lignes des correspondances.

$ git grep -n gmtime\_r

La commande grep peut etre enrichie grace a un certain nombre d’options nteressantes.

vous pouvez indiquer a Git de resumer le resultat en ne montrant que les fichiers et le nombre de correspondances au moyen de l’option --count :

$ git grep --count gmtime\_r

Si vous souhaitez voir dans quelle methode ou fonction la correspondance a ete trouvee, vous pouvez passer l’option -p :

$ git grep -p gmtime\_r \*.c

Vous pouvez aussi rechercher des combinaisons plus complexes de chaines de caracteres avec l’option --and

Ici, nous allons utiliser les options --break et --heading qui aident a decouper le resultat dans un format plus digeste.

Par exemple, recherchons toutes les lignes qui definissent une constante qui contient au choix ≪ LINK ≫ ou ≪ BUF\_MAX ≫ dans la base de code de Git avant la version 1.8.0.

$ git grep --break --heading -n -e '#define' --and ( -e LINK -e BUF\_MAX ) v1.8.0

La commande git grep a quelques avantages sur les commandes de recherche normales telles que grep et ack. Le premier est qu’elle est vraiment rapide, le second est qu’elle vous permet de rechercher dans n’importe quelle arborescence Git, pas seulement la copie de travail. Comme nous l’avons vu dans l’exemple ci-dessus, nous avons cherche des termes dans une version ancienne du code source de Git, pas dans la derniere version extraite.

**Recherche dans le journal Git**

Peut-etre ne cherchez-vous pas **ou** un terme apparait, mais plutot **quand** il existait ou fut introduit. La commande git log comprend un certain nombre d’outils puissants pour trouver des *commits* specifiques par le contenu de leurs messages ou le contenu des diffs qu’ils introduisent.

Si vous voulez trouver par exemple quand la constante ZLIB\_BUF\_MAX a ete initialement introduite, nous pouvons indiquez a Git de ne montrer que les *commits* qui soit ajoutent soit retirent la chaine avec l’option -S.

$ git log -SZLIB\_BUF\_MAX –oneline

Si vous devez etre plus specifique, vous pouvez fournir une expression reguliere a rechercher avec l’option -G.

**Recherche des evolutions d’une ligne**

On l’active avec l’option -L de git log et elle vous montre l’historique d’une fonction ou d’une ligne dans votre base de code.

Par exemple, si nous souhaitions voir toutes les modifications realisees sur la fonction git\_deflate\_bound dans le fichier zlib.c, nous pourrions lancer git log -L

:git\_deflate\_bound:zlib.c. Cette commande va essayer de determiner les limites de cette fonction et de rechercher dans l’historique chaque modification realisee sur la fonction comme une serie de patchs jusqu’au moment de sa creation.

$ git log -L :git\_deflate\_bound:zlib.c

Si Git ne peut pas determiner comment trouver la fonction ou la methode dans votre langage de programmation, vous pouvez aussi fournir une regex. Par exemple, cela aurait donne le meme resultat avec

git log -L '/unsigned long git\_deflate\_bound/',/^}/:zlib.c. Vous auriez pu aussi specifier un intervalle de lignes ou un numero de ligne et vous auriez obtenu le meme type de resultat.

**Reecrire l’historique**

**Modifier la derniere validation**

$ git commit –amend

Vous indexez les modifications que vous voulez en executant git add ou git rm, et le prochain git commit –amend prendra votre index courant et en fera le contenu de votre nouvelle validation.

* **Modifier plusieurs messages de validation**

Par exemple, si vous voulez modifier les 3 derniers messages de validation ou n’importe lequel des messages dans ce groupe, vous fournissez a git rebase -i le parent du dernier *commit* que vous voulez editer,

$ git rebase -i HEAD~3

Souvenez-vous egalement que ceci est une commande de rebasage, chaque *commit* inclus dans l intervalle HEAD~3..HEAD sera reecrit,

* **Reordonner les** *commits*
* **Ecraser un** *commit*
* **Diviser un** *commit*

Pour diviser un *commit*, il doit etre defait, puis partiellement indexe et valide autant de fois que vous voulez pour en finir avec lui

vous pouvez annuler (**reset**) ce *commit* et revalider les modifications que vous voulez pour creer plusieurs *commits*

$ git reset HEAD^

$ git add README

$ git commit -m 'updated README formatting'

$ git add lib/simplegit.rb

$ git commit -m 'added blame'

$ git rebase --continue

**L’option nucleaire :** filter-branch

Il existe une autre option de la reecriture d’historique que vous pouvez utiliser si vous avez besoin de reecrire un grand nombre de *commits* d’une maniere scriptable ; par exemple, modifier globalement votre adresse mail ou supprimer un fichier de tous les *commits*. La commande est filter-branch, et elle peut réécrire des pans entiers de votre historique, vous ne devriez donc pas l’utiliser a moins que votre projet ne soit pas encore public ou que personne n’ait encore travaille

sur les *commits* que vous allez reecrire. Cependant, cela peut etre tres utile. Vous allez maintenant apprendre quelques usages communs pour vous donner une idee de ses capacites

**Supprimer un fichier de chaque** *commit*

Cela arrive assez frequemment. Quelqu’un a accidentellement valide un enorme fichier binaire avec une commande git add . irreflechie, et vous voulez le supprimer partout. Vous avez peut-etre valide un fichier contenant un mot de passe et vous voulez rendre votre projet open source. Filterbranch est l’outil que vous voulez probablement utiliser pour nettoyer votre historique entier. Pour supprimer un fichier nomme ≪ passwords.txt ≫ de tout votre historique, vous pouvez utiliser l’option

--tree-filter de filter-branch :

$ git filter-branch --tree-filter 'rm -f passwords.txt' HEAD

Rewrite 6b9b3cf04e7c5686a9cb838c3f36a8cb6a0fc2bd (21/21)

Ref 'refs/heads/master' was rewritten

L’option --tree-filter execute la commande specifiee pour chaque *commit* et le revalide ensuite. Dans le cas present, vous supprimez le fichier nomme ≪ passwords.txt ≫ de chaque contenu, qu’il existait ou non. Si vous voulez supprimer tous les fichiers temporaires des editeurs validesaccidentellement, vous pouvez executer une commande telle que

git filter-branch --tree-filter "find \* -type f -name '\*~' -delete" HEAD.

Vous pourrez alors regarder Git reecrire l’arbre des *commits* et revalider a chaque fois, pour finir en modifiant la reference de la branche. C’est generalement une bonne idee de le faire dans un branche de test puis de faire une reinitialisation forte (**hard-reset**) de votre branche master si le resultat vous convient. Pour executer filter-branch sur toutes vos branches, vous pouvez ajouter --all a la commande.

**Reset demystifie**

Le moyen le plus simple de penser a reset et checkout consiste a representer Git comme un gestionnaire de contenu de trois arborescences differentes.

Git, comme systeme, gere et manipule trois arbres au cours de son operation normale :

**Arbre Role**

HEAD instantane de la derniere validation, prochain parent

Index instantane propose de la prochaine validation

Repertoire de travail bac a sable

**HEAD**

HEAD est un pointeur sur la reference de la branche actuelle, qui est a son tour un pointeur sur le dernier *commit* realise sur cette branche. Ceci signifie que HEAD sera le parent du prochain *commit* a creer.

$ git cat-file -p HEAD

$ git ls-tree -r HEAD

**L’index**

L’index est votre **prochain commit propose**. Nous avons aussi fait reference a ce concept comme la ≪ zone de preparation ≫ de Git du fait que c’est ce que Git examine lorsque vous lancez git commit.

$ git ls-files –s

Encore une fois, nous utilisons ici ls-files qui est plus une commande de coulisses qui vous montre l’etat actuel de votre index. L’index n’est pas techniquement parlant une structure arborescente ‑ c’est en fait un manifeste aplati ‑ mais pour nos besoins, c’est suffisamment proche.

**Le repertoire de travail**

Il faut penser a la copie de travail comme un **bac a sable** ou vous pouvez essayer vos modifications avant de les transferer dans votre index puis le valider dans votre historique.

Reset manipule directement les trois arborescences d’une manière simple et prédictible. Il réalise jusqu’à trois opérations basiques.

**Etape 1: deplacer HEAD**

La premiere chose que reset va faire consiste a deplacer ce qui est pointe par HEAD. Ce n’est pas la meme chose que changer HEAD lui-meme (ce que fait checkout). reset deplace la branche que HEAD pointe.

Quelle que soit la forme du reset que vous invoquez pour un *commit*, ce sera toujours la premiere chose qu’il tentera de faire. Avec reset --soft, il n’ira pas plus loin.

**Etape 2 : Mise a jour de l’index (--mixed)**

**Etape 3: Mise a jour de la copie de travail (--hard)**

La troisième chose que reset va faire est de faire correspondre la copie de travail avec l’index. Si vous utilisez l’option --hard, il continuera avec cette etape.

Il est important de noter que cette option (--hard) est le seul moyen de rendre la commande reset dangereuse et est un des tres rares cas ou Git va reellement detruire de la donnee. Toute autre invocation de reset peut etre defaite,

**Recapitulatif**

La commande reset remplace ces trois arbres dans un ordre specifique, s’arretant lorsque vous lui indiquez :

1. Deplace la branche pointee par HEAD *(s’arrete ici si* --soft*)*

2. Fait ressembler l’index a HEAD *(s’arrete ici a moins que* --hard*)*

3. Fait ressembler le repertoire de travail a l’index.

**Reset avec un chemin**

Donc, supposons que vous lancez git reset file.txt. Cette forme (puisque vous n’avez pas specifie un SHA-1 de commit ni de branche, et que vous n’avez pas non plus specifie --soft ou --hard) est un raccourcis pour git reset --mixed HEAD file.txt.

il ne fait que copier file.txt de HEAD vers index

Ceci a l’effet pratique de *desindexer* le fichier. Si on regarde cette commande dans le diagramme et qu’on pense a ce que git add fait, ce sont des opposes exacts

Nous pourrions tout aussi bien ne pas laisser Git considerer que nous voulions dire ≪ tirer les donnees depuis HEAD ≫ en specifiant un *commit* specifique d’ou tirer ce fichier. Nous lancerions juste quelque chose comme git reset eb43bf file.txt.

**Et checkout**

Comme reset, checkout manipule les trois arborescences et se comporte generalement differemment selon que vous indiquez un chemin vers un fichier ou non.

**Sans chemin**

Lancer git checkout [branche] est assez similaire a lancer git reset --hard [branche] en ce qu’il met a jour les trois arborescences pour qu’elles ressemblent a [branche], mais avec deux differences majeures

Premierement, a la difference de reset --hard, checkout preserve le repertoire de travail ; il s’assure de ne pas casser des fichiers qui ont change. En fait, il est meme un peu plus intelligent que ca – il essaie de faire une fusion simple dans le repertoire de travail, de facon que tous les fichiers **non modifies** soient mis a jour. reset --hard, par contre, va simplement tout remplacer unilateralement sans rien verifier.

Donc, dans les deux cas, nous deplacons HEAD pour pointer sur le commit A, mais la manière differe beaucoup. reset va deplacer la branche pointee par HEAD, alors que checkout va deplacer HEAD lui-meme.



**Avec des chemins**

L’autre facon de lancer checkout est avec un chemin de fichier, ce qui, comme reset, ne deplace pas HEAD. Cela correspond juste a git reset [branche] fichier car cela met a jour l’index avec ce fichier a ce *commit*, mais en remplacant le fichier dans le repertoire de travail. Ce serait exactement comme git reset --hard [branche] fichier (si reset le permettait) – cela ne preserve pas le repertoire de travail et ne deplace pas non plus HEAD.

**Resume**

Voici un aide-memoire sur ce que chaque commande affecte dans chaque arborescence. La colonne ≪ HEAD ≫ contient ≪ REF ≫ si cette commande deplace la reference (branche) pointee par HEAD, et ≪ HEAD ≫ si elle deplace HEAD lui-meme. Faites particulierement attention a la colonne ≪ preserve RT ? ≫ (preserve le repertoire de travail) – si elle indique **NON**, reflechissez a deux fois avant de lancer la commande.



**Fusion avancee**

**Conflits de fusion**

Premierement, nous entrons dans l’etat de conflit de fusion. Puis, nous voulons obtenir des copies de la version locale (**ours**), de la version distante (**theirs**, celle qui vient de la branche a fusionner) et de la version commune (l’ancetre commun depuis lequel les branches sont parties). Ensuite, nous voulons corriger au choix la version locale ou la distante et reessayer de fusionner juste ce fichier.

Obtenir les trois versions des fichiers est en fait assez facile. Git stocke toutes ces versions dans l’index sous formes d’etapes (**stages**), auxquelles chacune y a un nombre associe. Stage 1 est l’ancetre commun, stage 2 est notre version, stage 3 est la version de MERGE\_HEAD, la version qu’on cherche a fusionner (**theirs**).

Vous pouvez extraire une copie de chacune de ces versions du fichier en conflit avec la commande git show et une syntaxe speciale.

$ git show :1:hello.rb > hello.common.rb

$ git show :2:hello.rb > hello.ours.rb

$ git show :3:hello.rb > hello.theirs.rb

Si vous voulez rentrer un peu plus dans le dur, vous pouvez aussi utiliser la commande de plomberie ls-files -u pour recuperer les SHA-1 des blobs Git de chacun de ces fichiers.

$ git ls-files –u

La syntaxe :1:hello.rb est juste un raccourcis pour la recherche du SHA-1 de ce blob.

A present que nous avons le contenu des trois etapes dans notre repertoire de travail, nous pouvons reparer manuellement la copie distante pour resoudre le probleme d’espaces et refusionner le fichier avec la commande meconnue git merge-file dont c’est l’exacte fonction

$ git merge-file -p hello.ours.rb hello.common.rb hello.theirs.rb > hello.rb

vous pouvez demander a git diff de comparer le contenu de votre repertoire de travail que vous etes sur le point de valider comme resultat de la fusion avec n’importe quelle etape. Detaillons chaque comparaison. Pour comparer votre resultat avec ce que vous aviez dans votre branche avant la fusion, en d’autres termes, ce que la fusion a introduit, vous pouvez lancer git diff –ours

Si nous voulons voir le resultat de la fusion modifiee depuis la version distante, nous pouvons lancer git diff --theirs.

Enfin, nous pouvons voir comment le fichier a ete modifie dans les deux branches avec git diff --base.

**Examiner les conflits**

Si nous ouvrons le fichier, nous verrons quelque chose comme :

#! /usr/bin/env ruby

def hello

<<<<<<< HEAD

puts 'hola world'

======

puts 'hello mundo'

>>>>>>> mundo

end

hello()

Peut-etre le moyen de resoudre n’est-il pas evident. Il necessite plus de contexte.

Un outil utile est git checkout avec l’option --conflict. Il va re-extraire le fichier et remplacer les marqueurs de conflit. Cela peut etre utile si vous souhaitez eliminer les marqueurs et essayer de resoudre le conflit a nouveau.

Vous pouvez passer en parametre a --conflict, soit diff3 soit merge (le parametre par defaut). Si vous lui passez diff3, Git utilisera une version differente des marqueurs de conflit, vous fournissant non seulement les versions locales (*ours*) et distantes (*theirs*), mais aussi la version ≪ base ≫ integree pour vous fournir plus de contexte.

$ git checkout --conflict=diff3 hello.rb

Une fois que nous l’avons lance, le fichier ressemble a ceci :

#! /usr/bin/env ruby

def hello

<<<<<<< ours

puts 'hola world'

||||||| base

puts 'hello world'

======

puts 'hello mundo'

>>>>>>> theirs

end

hello()

Si vous appreciez ce format, vous pouvez le regler comme defaut pour les futur conflits de fusion en renseignant le parametre merge.conflictstyle avec diff3.

$ git config --global merge.conflictstyle diff3

La commande git checkout peut aussi accepter les options --ours et --theirs, qui peuvent servir de moyen rapide de choisir unilateralement une version ou une autre sans fusion

Cela peut etre particulierement utile pour les conflits de fichiers binaires ou vous ne pouvez que choisir un des cote, ou des conflits ou vous souhaitez fusionner certains fichiers depuis d’autres branches - vous pouvez fusionner, puis extraire certains fichiers depuis un cote ou un autre avant de valider le resultat.

**Journal de fusion**

Un autre outil utile pour la resolution de conflits de fusion est git log. Cela peut vous aider a obtenir du contexte ce qui a contribue aux conflits. Parcourir un petit morceau de l’historique pour se rappeler pourquoi deux lignes de developpement ont touche au meme endroit dans le code peut s’averer quelque fois tres utile.

Pour obtenir une liste complete de tous les *commits* uniques qui ont ete introduits dans chaque branche impliquee dans la fusion, nous pouvons utiliser la syntaxe ≪ triple point ≫ que nous avons apprise dans Triple point.

$ git log --oneline --left-right HEAD...MERGE\_HEAD

Neanmoins, nous pouvons simplifier encore plus ceci pour fournir beaucoup plus de contexte. Si nous ajoutons l’option --merge a git log, il n’affichera que les *commits* de part et d’autre de la fusion qui modifient un fichier presentant un conflit.

$ git log --oneline --left-right –merge

Si nous lancons cela avec l’option -p a la place, vous obtenez les diffs limites au fichier qui s’est retrouve en conflit. Cela peut s’averer **vraiment** utile pour vous donner le contexte necessaire a la comprehension de la raison d’un conflit et a sa resolution intelligente.

**Defaire des fusions**

**Correction des references**

git reset --hard HEAD~

Le defaut de cette approche est qu’elle re-ecrit l’historique, ce qui peut etre problematique avec un depot partage.

**Inverser le** *commit*

Si les deplacements des pointeurs de branche ne sont pas envisageables, Git vous donne encore l’option de creer un nouveau *commit* qui defait toutes les modifications d’un autre deja existant. Git appelle cette option une ≪ inversion ≫ (*revert*), et dans ce scenario particulier, vous l’invoqueriez comme ceci :

$ git revert -m 1 HEAD

**Autres types de fusions**

**Preference** *our* **ou** *theirs*

Premierement, il existe un autre mode utile que nous pouvons utiliser avec le mode ≪ recursive ≫ normal de fusion. Nous avons deja vu les options ignore-all-space et ignore-space-change qui sont passees avec -X mais nous pouvons aussi indiquer a Git de favoriser un cote plutot que l’autre lorsqu’il rencontre un conflit.

Si vous preferez que Git choisisse simplement un cote specifique et qu’il ignore l’autre cote au lieu de vous laisser fusionner manuellement le conflit, vous pouvez passer -Xours ou -Xtheirs a la commande merge.

Si une des options est specifiee, Git ne va pas ajouter de marqueurs de conflit. Toutes les differences qui peuvent etre fusionnees seront fusionnees. Pour toutes les differences qui generent un conflit, Git choisira simplement la version du cote que vous avez specifie, y compris pour les fichiers binaires.

$ git merge -Xours mundo

Cette option peut aussi etre passee a la commande git merge-file que nous avons utilisee plus tot en lancant quelque chose comme git merge-file --ours pour les fusions de fichiers individuels

Si vous voulez faire quelque chose similaire mais indiquer a Git de ne meme pas essayer de fusionner les modifications de l’autre cote, il existe une option draconienne qui s’appelle la *strategie* de fusion ≪ *ours* ≫.

Cela realisera une fusion factice. Cela enregistrera un nouveau *commit* de fusion avec les deux branches comme parents, mais ne regardera meme pas la branche en cours de fusion. Cela enregistrera simplement le code exact de la branche courante comme resultat de la fusion.

$ git merge -s ours mundo

Vous pouvez voir qu’il n’y a pas de difference entre la branche sur laquelle nous etions precedemment et le resultat de la fusion.

**Deboguer avec Git**

vous pouvez visualiser le fichier annote avec git blame pour voir quand chaque ligne de la methode a ete modifiee pour la derniere fois et par qui. Cet exemple utilise l’option -L pour limiter la sortie des lignes 12 a 22 :

$ git blame -L 12,22 simplegit.rb

**Sous-modules**

**Demarrer un sous-module**

Pour ajouter un nouveau sous-module, nous utilisons la commande git submodule add avec l’URL du projet que nous souhaitons suivre

$ git submodule add https://github.com/chaconinc/DbConnector

Si vous lancez git status a ce moment, vous noterez quelques differences.

Premierement, un fichier .gitmodules vient d’apparaitre. C’est le fichier de configuration qui stocke la liaison entre l’URL du projet et le sous-repertoire local dans lequel vous l’avez tire.

$ cat .gitmodules

[submodule "DbConnector"]

path = DbConnector

url = <https://github.com/chaconinc/DbConnector>

fichier. Il est important de noter que ce fichier est en gestion de version comme vos autres fichiers, a l’instar de votre fichier .gitignore. Il est pousse et tire comme le reste de votre projet.

$ git diff DbConnector

Meme si DbConnector est un sous-repertoire de votre repertoire de travail, Git le voit comme un sous-module et ne suit pas son contenu (si vous n’etes pas dans ce repertoire). En echange, Git l’enregistre comme un *commit* particulier de ce depot. Si vous souhaitez une sortie diff plus agreable, vous pouvez passer l’option --submodule a git diff.

$ git diff –submodule

**Cloner un projet avec des sous-modules**

Maintenant, vous allez apprendre a cloner un projet contenant des sous-modules. Quand vous recuperez un tel projet, vous obtenez les differents repertoires qui contiennent les sous-modules, mais encore aucun des fichiers

Le repertoire DbConnector est present mais vide. Vous devez executer deux commandes : git submodule init pour initialiser votre fichier local de configuration, et git submodule update pour tirer toutes les donnees de ce projet et recuperer le *commit* approprie tel que liste dans votre superprojet:

$ git submodule init

$ git submodule update

Il existe une autre maniere plus simple d’arriver au meme resultat. Si vous passez l’option --recurse-submodules a la commande git clone, celle-ci initialisera et mettra a jour automatiquement chaque sous-module du depot.

**Travailler sur un projet comprenant des sous-modules**

**Tirer des modifications amont**

Quand vous souhaitez vérifier si le sous-module a évolue, vous pouvez vous rendre dans le répertoire correspondant et lancer git fetch puis git merge de la branche amont pour mettre a jour votre code local.

$ git fetch

$ git merge origin/master

Si vous revenez maintenant dans le projet principal et lancez git diff --submodule, vous pouvez remarquer que le sous-module a ete mis a jour et vous pouvez obtenir une liste des *commits* qui y ont ete ajoutes

Il existe aussi un moyen plus facile, si vous preferez ne pas avoir a recuperer et fusionner manuellement les modifications dans le sous-repertoire. Si vous lancez la commande git submodule update --remote, Git se rendra dans vos sous-modules et realisera automatiquement le fetch et le merge.

C’est une information interessante car vous pouvez voir le journal des modifications que vous vous appretez a valider dans votre sous-module. Une fois validees, vous pouvez encore visualiser cette information en lancant git log -p.

$ git log -p –submodule

Par defaut, Git essaiera de mettre a jour **tous** les sous-modules lors d’une commande git submodule update --remote, donc si vous avez de nombreux sous-modules, il est preferable de specifier le sousmodule que vous souhaitez mettre a jour.

$ git submodule update --remote DbConnector

**Travailler sur un sous-module**

En premier, rendons-nous dans le repertoire de notre sous-module et extrayons une branche.

$ git checkout stable

Attaquons-nous au choix de politique de gestion. Pour le specifier manuellement, nous pouvons simplement ajouter l’option --merge or --rebasea l’appel de update.

$ git submodule update --remote –merge

Maintenant, si nous mettons a jour notre sous-module, nous pouvons voir ce qui arrive lors d’un rebasage de deux modifications concurrentes.

$ git submodule update --remote –rebase

**Publier les modifications dans un sous-module**

Pour etre sur que cela n’arrive pas, vous pouvez demander a Git de verifier que tous vos sousmodules ont ete correctement pousses avant de pouvoir pousser le projet principal. La commande git push accepte un argument --recurse-submodules qui peut avoir pour valeur ≪ *check* ≫ ou ≪ *on-demand* ≫. L’option ≪ check ≫ fera echouer push si au moins une des modifications des sous-modules n’a pas ete poussee.

Comme vous pouvez le voir, Git s’est rendu dans le module DbConnector et l’a pousse avant de pousser le projet principal. Si la poussee du sous-module echoue pour une raison quelconque, la poussee du projet principal sera annulee.

**Trucs et astuces pour les sous-modules**

Il existe quelques commandes qui permettent de travailler plus facilement avec les sous-modules.

**Submodule foreach**

Il existe une commande submodule foreach qui permet de lancer une commande arbitraire dans chaque sous-module. C’est particulierement utile si vous avez plusieurs sous-modules dans le meme projet

$ git submodule foreach git stash

Ensuite, nous pouvons creer une nouvelle branche et y basculer dans tous nos sous-modules.

$ git submodule foreach 'git checkout -b featureA'

$ git diff; git submodule foreach 'git diff'

Les alias Git, mais voici un exemple d’alias que vous pourriez trouver utiles si vous voulez travailler serieusement avec les sous-modules de Git.

$ git config alias.sdiff '!'"git diff && git submodule foreach 'git diff'"

$ git config alias.spush 'push --recurse-submodules=on-demand'

$ git config alias.supdate 'submodule update --remote --merge'

**Github**

git ls-remote origin affcihent la liste de toutes les branches et etiquettes ainsi que d’autres references dans le depot.

**Git dans d’autres environnements**

gitk est l’outil de visualisation graphique d’historique. Voyez-le comme une interface GUI puissante par-dessus git log et git grep. C’est l’outil a utiliser lorsque vous essayez de trouver un evenement passe ou de visualiser l’historique de votre projet.

Gitk accepte de nombreuses options de ligne de commande, dont la plupart sont passees directement a la commande git log sous-jacente. L’une des plus interessantes est probablement d’ajouter l’option --all qui indique a gitk de montrer tous les *commits* joignables depuis *n’importe quelle* reference, et pas seulement HEAD.

git-gui, par contre est un outil permettant de ciseler les *commits*. Lui aussi est plus facile a invoquer en ligne de commande :

$ git gui

**Commandes Git**

**Git distribue**

$ git checkout -b fonctionBv2 origin/master

$ git merge --no-commit --squash fonctionB

# (changement d'implementation)

$ git commit

$ git push macopie fonctionBv2

L’option --squash prend tout le travail de la branche a fusionner et le colle dans un *commit* sans fusion au sommet de la branche extraite. L’option --no-commit indique a Git de ne pas enregistrer automatiquement une validation. Cela permet de reporter toutes les modifications d’une autre branche, puis de realiser d’autres modifications avant de realiser une nouvelle validation

**Determiner les modifications introduites**

Faire une revue de tous les *commits* dans cette branche s’avere souvent d’une grande aide. Vous pouvez exclure les *commits* de la branche master en ajoutant l’option --not devant le nom de la branche. C’est equivalent au format master..contrib utilise plus haut. Par exemple, si votre contributeur vous envoie deux patchs et que vous creez une branche appelee contrib et y appliquez ces patchs, vous pouvez lancer ceci :

$ git log contrib --not master

Pour visualiser les modifications que chaque **commit** introduit, souvenez-vous que vous pouvez passer l’option -p a git log et elle ajoutera le diff introduit a chaque *commit*.

Pour visualiser un diff complet de ce qui arriverait si vous fusionniez cette branche thematique avec une autre branche, vous pouvez utiliser un truc bizarre pour obtenir les resultats corrects. Vous pourriez penser a lancer ceci :

$ git diff master

Si master est un ancetre directe de la branche thematique, ce n’est pas un probleme. Si les deux historiques ont diverge, le diff donnera l’impression que vous ajoutez toutes les nouveautes de la branche thematique et retirez tout ce qui a ete fait depuis dans la branche master.

Ce que vous souhaitez voir en fait, ce sont les modifications ajoutees sur la branche thematique — le travail que vous introduirez si vous fusionnez cette branche dans master. Vous obtenez ce resultat en demandant a Git de comparer le dernier instantane de la branche thematique avec son ancetre commun a la branche master le plus recent.

Techniquement, c’est realisable en determinant exactement l’ancetre commun et en lancant lacommande diff dessus :

$ git merge-base contrib master

36c7dba2c95e6bbb78df

$ git diff 36c7db

Neanmoins, comme ce n’est pas tres commode, Git fournit un raccourci pour realiser la meme chose : la syntaxe a trois points. Dans le contexte de la commande diff, vous pouvez placer trois points apres une autre branche pour realiser un diff entre le dernier instantane de la branche sur laquelle vous vous trouvez et son ancetre commun avec une autre branche :

$ git diff master...contrib

Cette commande ne vous montre que les modifications que votre branche thematique a introduites depuis son ancetre commun avec master. C’est une syntaxe tres simple a retenir

**Rerere**

Si vous fusionnez et rebasez beaucoup ou si vous maintenez une branche au long cours, la fonctionnalite appelee ≪ rerere ≫ peut s’averer utile.

Rerere signifie ≪ *re* utiliser les *re* solutions en *re* gistrees ≫ (“ *reuse recorded resolution* ”) ‑ c’est un moyen de raccourcir les resolutions manuelles de conflit. Quand rerere est actif, Git va conserver un jeu de couples d’images pre et post fusion des fichiers ayant presente des conflits, puis s’il s’apercoit qu’un conflit ressemble a une de ces resolutions, il va utiliser la meme strategie sans rien

vous demander.

**Generation d’un nom de revision**

Comme Git ne fournit pas par nature de nombres croissants tels que ≪ r123 ≫ a chaque validation, la commande git describe permet de generer un nom humainement lisible pour chaque *commit*. Git concatene le nom de l’etiquette la plus proche, le nombre de validations depuis cette etiquette et un code SHA-1 partiel du *commit* que l’on cherche a definir :

$ git describe master

v1.6.2-rc1-20-g8c5b85c

La commande git describe repose sur les etiquettes annotees (etiquettes creees avec les options –a ou -s). Les etiquettes de publication doivent donc etre creees de cette maniere si vous souhaitez utiliser git describe pour garantir que les *commits* seront decrits correctement. Vous pouvez aussi utiliser ces noms comme cible lors d’une extraction ou d’une commande show, bien qu’ils reposent sur le SHA-1 abrege et pourraient ne pas rester valides indefiniment.

**Preparation d’une publication**

Une des etapes consiste a creer une archive du dernier instantane de votre code pour les malheureux qui n’utilisent pas Git. La commande dediee a cette action est git archive :

$ git archive master --prefix='projet/' | gzip > `git describe master`.tar.gz

$ ls \*.tar.gz

v1.6.2-rc1-20-g8c5b85c.tar.gz

$ git archive master --prefix='project/' --format=zip > `git describe master`.zip

**Shortlog**

Il est temps d’envoyer une annonce a la liste de diffusion des nouveautes de votre projet. Une maniere simple d’obtenir rapidement une sorte de liste des modifications depuis votre derniere version ou courriel est d’utiliser la commande git shortlog. Elle resume toutes les validations dans l’intervalle que vous lui specifiez. Par exemple, ce qui suit vous donne un resume de toutes les validations depuis votre derniere version si celle-ci se nomme v1.0.1 :

$ git shortlog --no-merges master --not v1.0.1